P–Pot Stable and Super–Heavy–Tailed Distributions

R.-D. Reiss and U. Cormann

University of Siegen

November 2007

Research Group SRC, University of Siegen

Statistics, Riskanalysis, Computing (SRC)

Prof. Dr. R.-D. Reiss

Prof. Dr. H.-P. Scheffler

N.N.

Prof. Dr. E. Kaufmann

Dipl. Wirt-Math. P. Schupp

M.Sc. M. Frick

M.Sc. A. Kehl

M.Sc. Manjunath B.G.

M.Sc. U. Cormann

Dipl. Math. I. Schiltz

Contents

Extreme Value Distributions (EVDs) I

The df of the maximum of iid random variables $X_1, \ldots, X_n$ with common df F is given by

$$P\Big\{\max\{X_1, \ldots, X_n\} \le x\Big\} = F^n(x).$$

Max-Stability: A df F is max-stable if

$$F^n(d_n + c_n x) = F(x).$$

The max–stable dfs constitute the parametric family of EVDs with shape parameter α (Gumbel, Frechet, Weibull dfs).

Gumbel:	$G_0(x) = \exp(-e^{-x})$,	for all x;
Fréchet, $\alpha > 0$:	$G_{1,\alpha}(x) = \exp(-x^{-\alpha})$,	$x \ge 0$,
Weibull, $\alpha < 0$:	$G_{2,\alpha}(x) = \exp(-(-x)^{-\alpha})$,	$x \le 0$.

Extreme Value Distributions (EVDs) II

One can construct a continuous, parametric family of extreme value dfs G_γ with shape parameter $\gamma = 1/\alpha$. We get EVDs

$$G_\gamma(x) = \exp\left(-(1+\gamma x)^{-1/\gamma}\right).$$

in the γ–parametrization (von Mises–Jenkinson parametrization).

Generalized Pareto Distributions (GPDs) I

The exceedance df at the threshold u of a random variable X with df F is given by

$$P(X \le x | X > u) = \frac{F(x) - F(u)}{1 - F(u)} = F^{[u]}(x) .$$

POT-Stability: A df F is pot-stable if

$$F^{[u]}(b_u + a_u x) = F(x).$$

The possible pot–stable dfs constitute the parametric family of GPDs with shape parameter α (exponential, Pareto, beta dfs).

Exponential:	$W_0(x) = 1 - e^{-x},$	$x \ge 0,$
Pareto, $\alpha > 0$:	$W_{1,\alpha}(x) = 1 - x^{-\alpha},$	$x \ge 1,$
Beta, $\alpha < 0$:	$W_{2,\alpha}(x) = 1 - (-x)^{-\alpha},$	$-1 \le x \le 0.$

We have $W = 1 + \log G, \quad \text{if } \log G > -1.$

Generalized Pareto Distributions (GPDs) II

One can construct a continuous, parametric family of GPDs W_γ with shape parameter $\gamma = 1/\alpha$ by

$$W_\gamma(x) = 1 - (1 + \gamma x)^{-1/\gamma}, \quad x \geq 0,$$

which is the von Mises–Jenkinson parametrization.

Heavy–Tailed Distributions

- E.g. Frechet, Pareto and sum–stable (except of Gaussian) distributions are heavy–tailed in the upper tail in so far that certain moments are infinite,
- the variance (respectively, the expectation) of an EV, Pareto and sum–stable (except of Gaussian) random variable is infinite if $\alpha \leq 2$ (respectively, $\alpha \leq 1$),
- in the literature, data are frequently classified as heavy–tailed with $\alpha \leq 2$ due to the wrong choice of estimator.

Pareto as Log-Exponential Model

We start with a rv X with exponential df

$$H(x) = 1 - \exp(-x), \quad x \geq 0.$$

We include a scale parameter $\gamma > 0$. Then,

- $Y = \exp(\gamma X) - 1$ has the df

$$\widetilde{W}_{\gamma}(x) = 1 - (1 + x)^{-1/\gamma}, \quad x \geq 0,$$

- adding a scale parameter $\sigma > 0$ one gets the log–exponential df

$$\widetilde{W}_{\gamma,\sigma}(x) = 1 - (1 + x/\sigma)^{-1/\gamma}, \quad x \geq 0,$$

  which is a Pareto df with shape parameter $\gamma > 0$ and scale parameter $\sigma > 0$.

The Role of the Scale Parameter

- if the scale parameter σ is not added in the Pareto model, then carry out statistical procedures based on data $\log(1+y_i)$ in the original exponential model; e.g., use the MLE for the scale parameter γ in the exponential model,
- the scale parameter σ in the Pareto model may be regarded as a second–order–parameter in so far that it "vanishes" if the threshold u goes to ∞, yet the influence can be disastrous for finite thresholds u,
- one may repair statistical procedures defined for the reduced model (that is, without the scale parameter) under second–order–conditions; e.g. the Hill estimator can be repaired by bias–reduction.

Generalized Pareto as Generalized Log-Exponential Model

Varying the former scale parameter γ one gets

$$W_{\gamma,\sigma} := \widetilde{W}_{\gamma,\sigma/\gamma}(x) \to_{\gamma\to 0} H(x/\sigma) = W_{0,\sigma}(x).$$

With

$$\{W_{\gamma,\sigma} : \gamma \text{ real}, \sigma > 0\}$$

we receive the family of generalized Pareto dfs (GPDs) in the von Mises–Jenkinson parametrization.

General Approach to Log-Families I

We formalize the preceeding procedure which led from exponential to Pareto dfs.

- Let H_ϑ be a family of dfs, where ϑ is a shape parameter and let X have the df H_ϑ.
- We include a scale parameter $\beta > 0$ in the original model: then,

$$Y = \exp(\beta X) - 1$$

  has the df $H_\vartheta\left(\frac{1}{\beta}\log(1+x)\right)$.
- Adding a scale parameter σ one obtains the full log–family

$$F_{\vartheta,\beta,\sigma}(x) = H_\vartheta\left(\frac{1}{\beta}\log(1+x/\sigma)\right) \tag{1}$$

  with shape parameters ϑ, β and scale parameter σ.

General Approach to Log-Families II

The original family of dfs H_ϑ is included in the log–family in the limit. Varying the former scale parameter β one gets

$$F_{\vartheta,\beta,\sigma/\beta}(x) \to H_\vartheta(x/\sigma), \quad \beta \to 0. \tag{2}$$

Genesis of Log–Pareto

Next we focus on log–Pareto dfs which are first of all exponential transformations of Pareto dfs. We mention four different approaches:

i. log–family approach
ii. mixtures of Pareto dfs (Reiss, Thomas (1997))
iii. p–pot stability, domains of attraction
iv. slowly varying approach (Alves et al. (2006), Meerschaert, Scheffler (2006))

Further Literature

Until recently log–Pareto distributions have rarely been studied in the statistical literature (e.g., they are not mentioned in the book by Johnson et. al (1994)). Notable exeptions are

i. Galambos (1987) and Subramanya (1994) as an example for dfs which are in no domain of attraction of EVDs with $F(x) = 1 - \frac{1}{\log(x)}$

ii. Desgagné, Angers (2005) in conjunction with generalized exponential power models

Log–Family–Approach

We apply the above mentioned general procedure:

- As mentioned before: starting with the exponential df $H(x/\gamma) = 1 - \exp(-x/\gamma)$, applying the exponential transformation and adding a scale parameter $\beta > 0$ one receives the Pareto df

$$\widetilde{W}_{\gamma,\beta}(x) = 1 - (1 + x/\beta)^{-1/\gamma}, \quad x > 0. \tag{3}$$

- Applying the exponential transformation to $\widetilde{W}_{\gamma,\beta}$ and adding a scale parameter σ, one gets the log–Pareto df

$$\tilde{L}_{\gamma,\beta,\sigma}(x) = 1 - \left(1 + \frac{1}{\beta}\log\left(1 + \frac{x}{\sigma}\right)\right)^{-1/\gamma}, \quad x > 0, \tag{4}$$

with two shape parameters γ, β and the scale parameter σ.

Super–Heavy Tailed Distributions

- All moments of log Pareto–dfs are infinite.
- For the log–moments one gets

$$\int (\log(1+x))^z \, d\tilde{L}_{\gamma,\beta,\sigma}(x) = \infty \qquad (5)$$

if $z \geq 1/\gamma$; that is, log–Pareto dfs possess super–heavy upper tails (Reiss, Thomas (1997)).

- Only the shape parameter γ is crucial for the existence of finite log–moments.

Pareto dfs as limits of log–Pareto dfs

The Pareto model can be interpreted as limiting case of the Log–Pareto model for varying shape parameters in two different manners.

i. Varying the former scale parameter β:

$$\tilde{L}_{\gamma,\beta,\sigma/\beta}(x) \to \widetilde{W}_{\gamma,\sigma}(x), \quad \beta \to 0,$$

ii. Because $W_{\gamma,\beta}(x) \to W_{0,\beta}$, $\gamma \to 0$, we also get

$$L_{\gamma,\beta,\sigma}(x) := \tilde{L}_{\gamma,\beta/\gamma,\sigma}(x) \to \widetilde{W}_{\beta,\sigma}(x) =: L_{0,\beta,\sigma}(x), \quad \gamma \to 0,$$

by varying the 2–step former scale parameter γ.

Generalized Log–Pareto Families

The family of log–Pareto distributions can be extended to the family of generalized log–Pareto distributions (GLPDs) if one includes first the shape parameter $\gamma = 0$ and, then also negative shape parameters γ.

- For real γ one obtains GLPDs of the form

$$L_{\gamma,\beta,\sigma}(x) = 1 - \left(1 + \frac{\gamma}{\beta} \log\left(1 + \frac{x}{\sigma}\right)\right)^{-1/\gamma} \tag{6}$$

  for $x > 0$ if $\gamma > 0$ and $0 < x < (\exp(\beta/|\gamma|) - 1)\,\sigma$ if $\gamma < 0$.

- For $\gamma = 0$ we have

$$L_{0,\beta,\sigma}(x) = \widetilde{W}_{\beta,\sigma}(x), \quad x > 0. \tag{7}$$

Iterated Log–Pareto Families

The log–approach can be further iterated.

- Applying the exponential transformation (and adding a scale parameter) to log–Pareto dfs one receives a second order log–Pareto df with three shape and one scale parameter, namely

$$L^{(2)}_{\gamma,\beta,\xi,\sigma}(x) = 1 - \left(1 + \frac{1}{\beta}\log\left(1 + \frac{1}{\xi}\log\left(1 + x/\sigma\right)\right)\right)^{-1/\gamma}. \qquad (8)$$

- This procedure can be iterated further on leading to dfs with more and more shape–parameters and higher order iterated heavy–tailed dfs.

P-Max Stability

P-Max-Stability: A df F is p-max-stable (Pancheva (1985)), if

$$F^n(\text{sign}(x)c_n|x|^{d_n}) = F(x).$$

The p-max-stable dfs can be derived from the family of extreme value distributions. We have for a p-max-stable df H with mass on the positive half–line:

$$H(x) = H_{\gamma,\beta}(x) = G_\gamma\left(\frac{1}{\beta}\log(x)\right), \quad \gamma \text{ real}, \beta > 0.$$

Thus we have a parametric family with two shape parameters.

P-POT Stability I

P-POT-Stability: A df F is p–pot–stable if

$$F^{[u]}(\mathrm{sign}(x) b_u |x|^{a_u}) = F(x).$$

The possible p–pot–stable dfs with mass on the positive half–line correspond to the parametric family of GLPDs with shape parameters γ and β. Let

$$\hat{L}_{\gamma,\beta}(x) = W_\gamma\left(\frac{1}{\beta}\log(x)\right).$$

We have $\hat{L} = 1 + \log H, \quad \text{if } \log H > -1.$

P–Pot Stability II

- The log–Pareto df $\hat{L}_{\gamma,\beta,\sigma}$ is p–pot stable in the sense that there exists β_u and α_u such that

$$\hat{L}^{[u]}_{\gamma,\beta,\sigma}\left(\alpha_u x^{\beta_u}\right) = \hat{L}_{\gamma,\beta,\sigma}(x), \quad x > 0. \tag{9}$$

- Choose $\alpha_u = u\sigma^{-\beta_u}$ and $\beta_u = 1 + \gamma/\beta \log(u/\sigma)$ if $u > \sigma$ and $\beta_u > 0$.

Limiting P–POT–DFS

Let F and L be dfs such that

$$F^{[u]}\left(\operatorname{sign}(x)\alpha(u)\,|x|^{\beta(u)}\right) \xrightarrow[u\to\omega(F)]{} L(x). \qquad (10)$$

We say that a df F is in the p–pot domain of attraction $\mathcal{D}_{p-pot}(L)$ of L [in short, $F \in \mathcal{D}_{p-pot}(L)$] if (10) holds.

If $\omega(F) > 0$, then GLPDs of the form

$$\hat{L}_{\gamma,\beta,\sigma}(x) = 1 - \left(1 + \frac{\gamma}{\beta}\log\left(x/\sigma\right)\right)^{-1/\gamma}, \quad \beta,\sigma > 0, \gamma \in \mathbb{R} \qquad (11)$$

are the only possible continuous limiting dfs in (10). Thus, GLPDs form a unified model for exceedances over high thresholds.

Goodness–of–fit Tests for Pareto Families

Goodness–of–fit tests for Pareto models are well known in the statistical literature, e.g., Hüsler and Li (2006) or Villasenõr et al. (2007).

One is testing the null hypothesis

$\mathcal{H}_0$: F is a Pareto df with parameter $\gamma \in \mathbb{R}$ and $\sigma > 0$

against the alternative

$\mathcal{H}_1$: F is not a Pareto df.

Goodness–of–fit Tests for Log–Pareto Models

Goodness–of–fit tests for log–Pareto models can be derived from the corresponding ones for Pareto models.

- A random variable X is distributed according to a log–Pareto df if and only if $Y = \log(1 + X/\sigma)$ is distributed according to a Pareto df.
- A goodness–of–fit test for a log–Pareto df based on a sample $x_1, \ldots, x_n$ can be established in the following manner:
  i. Find an estimator $\hat{\sigma}$ of the unknown scale parameter σ.
  ii. Consider the transformed data
  $$y_i = \log(1 + x_i/\hat{\sigma}) . \tag{12}$$
  iii. Finally, a goodness–of–fit test is carried out for the Pareto model based on the transformed data $y_1, \ldots, y_n$.

Final Remarks

Further research work:

- compare the different approaches
- identify second–order–parameters (scale as well as shape parameters)
- data analysis:
  - what are the appropriate tools?
  - do there exist relevant data sets?
- extend concepts from the univariate to the multivariate framework
- further statistical procedures
- define max– and pot–stability for other groups of transformations
- applications in finance and insurance and other fields

Literature I

- Christoph, G., Falk, M., "A note on domains of attraction of p–max stable laws," Statist. Probab. Letters 28, 279–284, (1994).
- Cormann, U. and Reiss, R.–D., "Generalizing the Pareto to the Log–Pareto model and statistical inference".
- Desgagné, A. and Angers, J.–F., "Importance sampling with the generalized exponential power density," Statist. Comp. 15, 189–195, (2005).
- Diebolt, J., El–Aroui, M.–A., Garrido, M. and Girard, S., "Quasi–conjugate Bayes estimates for GPD parameters and application to heavy tails modelling," available at: http://www.inria.fr/rrt/rr-4803.html, (2003).

Literature II

- Diebolt, J., El–Aroui, M.–A., Garrido, M. and Girard, S., “Quasi–conjugate Bayes estimates for GPD parameters and application to heavy tails modelling,” Extremes 8, 57–78, (2005).
- Falk, M., Hüsler, J. and Reiss, R.–D., Laws of Small Numbers: Extremes and Rare Events, DMV–Seminar Bd 23, Birkhäuser, Basel, 1994 (2nd ed. 2004).
- Fraga Alves, M.I., de Haan, L. and Neves, C., ‘Statistical inference for heavy and super–heavy tailed distributions,” available at: http://people.few.eur.nl/ldehaan/, (2006).
- Galambos, J., The Asymptotic Theory of Extreme Order Statistics, 2nd edition, Krieger, Malabar, 1987.
- Hüsler, J., Li, D., “On testing extreme value conditions,” Extremes 9, 69–86, (2006).

Literature III

- Johnson, N., Kotz, S. and Balakrishnan, N., Continuous Univariate Distributions, vol. 1-2. New York: John Wiley & Sons, 1994.
- Meerschaert, M.M. and Scheffler, H.-P., “Stochastic model for ultraslow diffusion”, Stoch. Process. Appl. 116, 1215–1235 (2006).
- Mohan, N.R. and Ravi, S., “Max domains of attraction of univariate and multivariate p–max stable laws”, Theory of Probability and its Applications 37, 632–643 (1993).
- Pancheva, E.I, “Limit theorems for extreme order statistics under non–linear transformations,” in: Stability Problems for Stochastic Models, V.V. Kalashnikov and V.M. Zolotarev eds., Lecture Notes in Statistics 1155, 248–309, Springer, Berlin 1985.
- Reiss, R.–D. and Cormann, U., An example of real–life data where the Hill estimator is correct, in: Advances in Mathematical and Statistical Modeling, B.C. Arnold et al. eds., Birkhäuser, Boston, to appear (2007)

Literature IV

- Reiss, R.–D. and Thomas, M., Statistical Analysis of Extreme Values, Birkhäuser, Basel 1997 (2nd ed. 2001, 3rd ed. 2007).
- Subramanya, U.R., On max domains of attraction of univariate p–max stable laws, Statist. Probab. Letters, 19, 271–279, (1994).
- Villasenõr–Alva, J.A., Gonzalez–Estrada, E. and Reiss, R.–D., “A nonparametric goodness of fit test for the generalized Pareto distribution,” in preparation.